VALUE ADDED PRODUCTS FROM MODIS TIME-SERIES DATA SETS TO SUPPORT DOI / USGS INVASIVE SPECIES MANAGEMENT

Jeffrey T. Morisette, Jeffrey A. Pedelty, and John L. Schnase
NASA Goddard Space Flight Center

Thomas J. Stohlgren
National Institute of Invasive Species Science
USGS Fort Collins Science Center

Abstract

We propose to develop a suite of MODIS time-series data products specifically tailored to invasive species management and policy decision-making for the Department of Interior's US Geological Survey. Recent work substantiates the importance of MODIS time-series data in geostatistical distribution models of invasive species. We will refine these results to produce image layers that summarize MODIS time-series data. These new summary layers will be used as explanatory variables within the USGS invasive species modeling framework and will feed directly into the National Invasive Species Forecasting System, which is a core component of an invasive species decision support capability being developed jointly by the USGS and NASA The specific MODIS land product being used as well as the techniques used to summarize the time-series will be based on collaboration with ecologists familiar with a given site or exotic species. The result will be a set of products designed to meet the specific needs of a large and expanding community of ecologists and natural resource managers who deal with invasive species. The work will contribute to NASA's goal of expanded use of MODIS by engaging a new client community that has a focused, thematic, and interdisciplinary data requirement of national importance. The new data products will be made available through the National Biological Information Infrastructure and the USGS National Institute of Invasive Species Science. These new data products are vital to understanding and managing the effects of invasive species on human health, the economy, native biodiversity, and ecosystem processes.

PROJECT DESCRIPTION

We are working with the NASA Office of Earth Science and the US Geological Survey to develop a National Invasive Species Forecasting System. These are the first steps toward a vision of a national decision support capability for the management and control of invasive species, perhaps the most devastating biological phenomenon of modern times. Building such a capability will require groundbreaking advances across a range of science and technology fronts and a scaling up of our ability to measure, model, and predict complex, interdependent physical and biological processes. A critical issue is the problem of developing new data products tailored to the specific needs of the invasive species research community. In particular, the lack of easily accessible, time-series data products is becoming a fundamental bottleneck in developing new applications and pursuing new science questions in this area. We recognize the possibility that others will propose to provide global phenology (or other time-series-related) MODIS products. However, we feel strongly that the "state of the science" in this arena is such that one-size-does-not-fit all. Here we propose a relatively low cost, low risk effort to explore methods of summarizing MODIS time-series data to support ecological forecasting within an existing system, at specific locations, for specific phenomenon.

We propose to address the issue of how best to exploit the MODIS time-series for invasive species forecasting by creating a suite of value added MODIS time-series data products specifically tailored to management and policy decision-making within the Department of Interior's US Geological Survey. The derived products will be designed to feed directly into the National Invasive Species Forecasting System and related applications. Over the next three years, we believe these new products will significantly improve our predictive capabilities and will become vital to understanding and managing the effects of invasive species on human health, the economy, native biodiversity, and ecosystem processes.

Background

Non-indigenous invasive species may pose the 21st century's single most formidable threat of natural disaster (NRC, 2000). During the past hundred years, non-indigenous plants, animals, and pathogens have been introduced at increasing rates into all US ecosystems. A growing number of these species are becoming invasive, and they contribute to declines in native species and to changes in ecosystem function. The direct cost to the American economy alone is estimated at $100-200 billion per year, greater than all other natural disasters combined (NISC, 2001).

An invasive species is defined as a non-native species whose introduction causes or is likely to cause harm to the economy, environment, or human health. The cost of infestations of leafy spurge alone to agricultural producers and taxpayers is $144 million a year in the Dakotas, Montana, and Wyoming. Aggressive invasive fishes in the Great Lakes threaten a commercial fishery valued at $4.5 billion that supports 81,000 jobs. Invasive Norway rats cause up to $19 billion in environmental and economic damage every year. Non-native livestock diseases cost $9 billion a year. In the coming decades, increasing human travel and trade and changing types and patterns of environmental disturbance are expected to exacerbate these impacts. Because of its high diversity of environmental conditions and habitats, the US is particularly vulnerable to invasions by non-indigenous species. (See *www.invasivespecies.gov* for additional information.)

The US Federal Government has begun to mount an organized effort to address the invasive species threat, coalescing around Executive Order 13112 (1999). The newly formed National Invasive Species Council has issued a draft National Invasive Species Management Plan and has assembled several technical working groups. The National Biological Information Infrastructure (NBII) has several regional programs developing invasive species information systems as their highest priority initiative and established a national node for invasive species in 2002. But the invasive species threat is not affecting the United States alone. Globalization has greatly increased the international movements of harmful species through travel and agricultural, horticultural, and pet industries; the invasive species threat has become a principal impediment to international trade agreements. So the US Federal Government's efforts are coordinated with international initiatives under the United Nations' Global Invasive Species Programme, the North American Free Trade Agreement's North American Biodiversity Information Network, the Summit of the Americas' Inter-American Biodiversity Information Network, and a number of bilateral agreements to develop international exchange on invasive species information.

In the United States, the USGS has a lead role in delivering scientific information concerning invasive species on Federal lands. USGS technical and scientific capabilities directly support the management of Department of Interior lands and waters by documenting, monitoring, and predicting the establishment and spread of invasive species. The USGS studies the ecology of invading species and vulnerable habitats to support prevention, early detection, assessment, containment, and, where possible, eradication of new invaders. The USGS also investigates the physical properties, composition, and hydrology of geologic substrates to identify lands vulnerable to invasion by exotic plants. USGS coordinates the interagency activities of NBII, whose extensive network of thematic and information service nodes play a critical role in delivering biological information of importance to invasive species.

NASA has a complementary and synergistic role to play in helping the USGS understand and manage invasive species. NASA's Earth Science Enterprise currently provides measurements from Terra, QuickSCAT, Landsat 7, Jason, and other missions that map key ecosystem attributes needed to predict invasive species distributions. A number of planned missions in the near- to mid-term will expand these measurements to include critical three-dimensional structure derived from synthetic aperture radar (SAR) and LIDAR technologies. Measurements are also supported through data purchase programs, including SeaWiFS ocean color imagery; high-resolution optical imagery from IKONOS, QuickBird, and other private-sector satellites; and land cover data from the Landsat Data Continuity Mission. In addition, NASA provides the computational capabilities and expertise in large-scale, coupled Earth system modeling needed to ensure the successful transfer of the system into operational use.

NASA's unique national role in developing remote sensing capabilities and in understanding the Earth system, coupled with the urgency of the invasive species threat, the ready partnership of the USGS, and the compelling societal and economic benefits of this effort, make these activities strategically important to NASA. However, success will ultimately require a far more comprehensive and sophisticated use of NASA data, particularly in the temporal domain.

That is the key problem that we hope to address in creating these new, value added MODIS time-series products.

The National Invasive Species Forecasting System

The USGS uses advanced geostatistical methods to integrate multiple types and scales of data, including satellite images, aerial photography, and ground data of various resolutions to map these resources (Stohlgren *et al.*, 1997a-c, 1999 a,b). The Invasive Species Forecasting System (ISFS) provides a framework for using USGS's early detection and monitoring protocols and predictive models to process NASA and commercial data and create on-demand, regional-scale assessments of invasive species patterns and vulnerable habitats. When fully implemented the forecasting system will be a dynamic and flexible mechanism for generating predictive maps of hot spots for potential exotic species invasions (Schnase et al., 2002b). These maps — when combined with economic, environmental, sociological, geographic, and other types of data and models — form a critical part of an invasive species decision support capability that will be used by private and public national, state, and local management agencies for remediation, management, and control.

Operational Context

Figure 1 shows the operational context for the Invasive Species Forecasting System. The ISFS will have users from government agencies, universities, industry, and the general public. To the users, the ISFS is deployed as a Web-based system that will present options for applying a series of models to available datasets yielding predictive result sets. The system will ingest data from different sources and in different formats, and existing models can be either run, or new ones created. The system will output maps and additional decision support information relating to predicted species distributions.

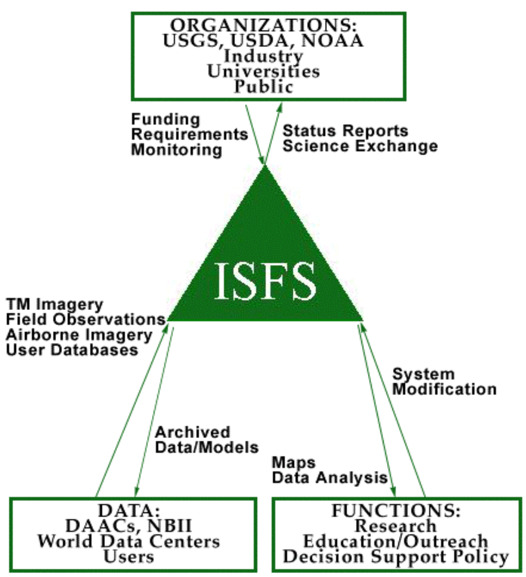


Figure 1: Context diagram for the Invasive Species Forecasting System. The proposed research would augment the data feeding into the lower left vertex of the ISFS.

Spatial Modeling

Mapping of biological resources is central to confronting the invasive species threat and forms the basis of the Invasive Species Forecasting System. The ability to model spatial variability in landscape characteristics requires the generation of full-coverage maps depicting characteristics measured in the field (Stohlgren *et al.*, 1997a-c, 1999a,b). Spatial statistics and geostatistics provide a means of developing spatial models that can be used to correlate coarse scale geographical data with micro-scale field measurements of biotic variables.

Figure 2 shows the USGS process for creating predictive spatial models. Basically, the process accepts as input a collection of ecological attributes, such as topographic data, species data, soil characteristics, ETM+ -derived vegetation indices, etc. These attributes are examined for statistically viable relationships between predictor variables and response variables. Trend surface analyses are performed, and residuals from the trend surface analyses are further analyzed for spatial structure using kriging and co-kriging. The results are brought together to produce a refined spatial prediction that is accompanied by an estimate of uncertainty. It is important to emphasize that the process's ability to produce both predictive maps and a maps of uncertainty significantly increases its value for decision support, since useful predictions are ultimately dependant upon a quantifiable understanding of error (OSTP, 2001; Smith *et al.*, 2001).

Figure 2: Diagram showing USGS geostatistical process for creating predictive spatial models for exotic plant species, uncertainty, forest parameters, soil variables, and other biotic and abiotic factors. The process relies on the creation of trend surface maps using regression, kriging, and co-kriging.

The USGS modeling framework thus combines multiple data layers, geographic information systems, image processing, and geostatistical techniques to identify patterns of biodiversity in geographic space. This is done by modeling the field-measured parameters (such as native or non-native species richness or diversity) with a set of independent variables (such as elevation, slope, aspect, soil properties, and spectral values from Landsat imagery). For example, one model for species richness could be written generally as:

$$Species\ Richness = f\left(DEM,\ soil\ layer,\ Landsat\ imagery\right)$$

where *f* is a linear function of variables derived from the input data sources. Within the ISFS, the function is fit using stepwise regression to identify the best combination of independent variables. Each independent variable in the model has a corresponding image or surface. The model fit with stepwise regression is then applied to the relevant input layers to produce a preliminary predicted surface. This linear model is further illustrated in Figure 3, which shows a representative set of input variables.

Standard linear regression techniques require the assumption of independent error terms. However, due to the spatial nature of the phenomenon being modeled, this assumption is seldom met. This issue is addressed by first analyzing the residuals resulting from the model for spatial autocorrelation by calculating both empirical variograms (Issaks and Srivastava, 1989; Deutsch and Journel, 1992) and the Moran's I test statistic (Haining, 1990). If these indicate spatial structure in the residuals, kriging is applied to the residuals to create a "residual surface" that is added to the preliminary predicted surface to produce the final predicted surface.

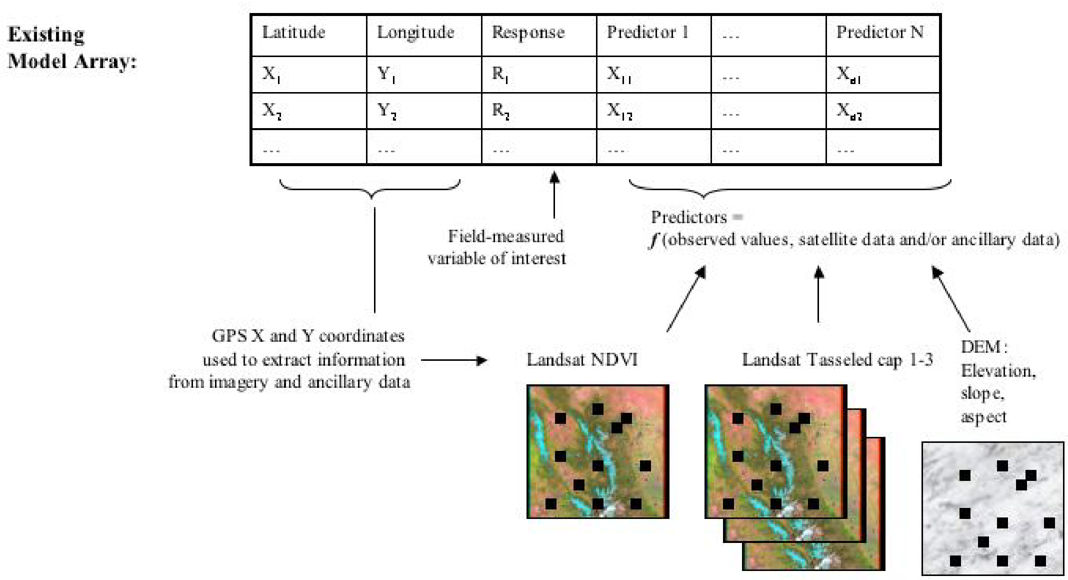


Figure 3: Diagram showing how the current linear model ingests remote sensing data and field data.

The models are cross-validated to assess variability in the prediction errors. Cross-validation includes deleting one or several observations from the data set and predicting the deleted observations using the remaining observations in the data set. This process is repeated for all observations in the data set. Finally, summary statistics of the estimated values are computed for the region.

USGS predictive models provide information on large-scale spatial variability that extends in crucial ways our understanding of the small-scale structure provided by field data. These techniques have been used successfully for determining potential distributions of exotic plants at the Cerro Grande Wildfire Site (shown in Figure 4), in Rocky Mountain National Park, and in Grand Staircase-Escalante National Monument. They are a core component in the decision support activities of the USGS and partner federal and state agencies. In order to scale these methods to larger applications, part of the ISFS effort is to improve the computational performance of the geostatistical algorithms and move the kriging calculation to cluster computers under a separate cooperative agreement with the NASA Computational Technologies Project. We have developed an initial parallel kriging code that gives greater than 93% scaling efficiency with 32 processors. This reduction in processing time will allow the USGS to extend the techniques to larger areas and to ingest more data layers. The next major opportunity to enable an operational Invasive Species Forecasting System is to increase the use of remotely sensed data, which is the major focus of the current proposal.

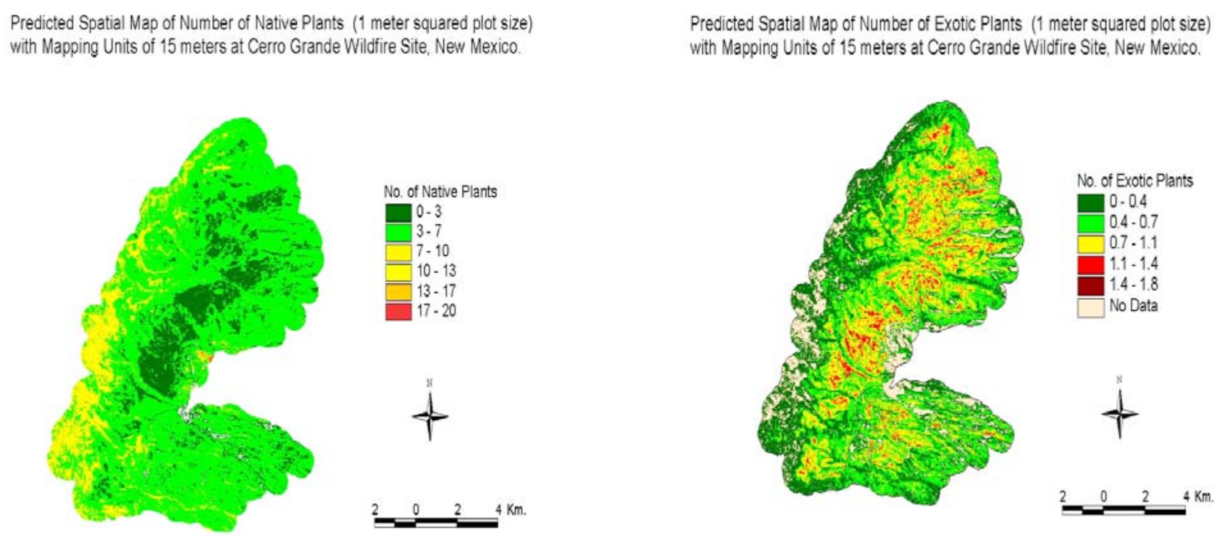


Figure 4: Predicted native and exotic species richness on the Cerro Grande Wildfire Site, Los Alamos, NM. This is an example of the type of predictive spatial map used by USGS in invasive species decision support.

Incorporating Time-Series Data

A vast literature points to the usefulness of time-series data in evaluating phenological patterns of biological importance. It is not surprising that significant improvements in model accuracy may be obtained by applying MODIS time-series data to problems dealing with invasive species.

Many non-native plant species enjoy the same environmental resources as native plant species: high light, soil nitrogen, water, and warm temperatures appear beneficial to plants regardless of origin (Stohlgren et al. 1999, 2001, 2003). However, surveys of plant invasions are usually limited to one-time measurements of environmental conditions such as plant foliar cover, soil chemistry, and soil texture. Time-series data allow scientists to quantify past, current, and potentially future resources available to invading species. For what are often rapidly proliferating invaders, it is particularly important to be able to use remote sensing information to determine the timing and amount of precipitation and green-up, factors that could be used to monitor plant species phenology (growth, seed-set) and water availability for range extensions.

Two particularly threatening examples are Tamarisk (salt cedar, *Tamarix sp.*), whose distributions are dependent on water availability in riparian zones, and Cheat Grass (*Bromus tectorum*), a non-native annual grass that responds strongly to early spring precipitation when many native species are still dormant. The timing and magnitude of Cheat Grass "green-up" can affect wildfire frequency and intensity in the summer – a major concern in many parts of the western US. Time-series and spatial modeling of environmental resources may greatly improve the spatial prediction and aid in the early detection of new invaders, or new populations of chronic pests.

Until recently, Landsat 7 Enhanced Thematic Mapper plus (ETM+) has been the primary satellite image data source used by USGS in modeling potential invasive species habitats and distribution. An advantage of this approach is the relatively high spatial resolution of the ETM+ sensor. However, there are disadvantages. ETM+ has a repeat cycle of only 16 days, and cloud cover can reduce the amount of useful imagery by lengthening the time between clear scenes to 32, 48, or more days. In addition, each scene costs $605 and requires approximately 600MB of computer file storage (for roughly 30,000km^2). Practical use of ETM+ data by land management agencies such as USGS is thus severely limited by cost, size, and the time intervals between cloud-free image acquisitions. Two or three scenes per year are not enough to distinguish critical phenological differences between native and exotic species.

In contrast, the MODIS instrument covers the globe with much higher temporal resolution. As shown in Figure 5, MODIS provides daily coverage for areas beyond +/- approximately 30^o, and MODIS imagery covers the entire globe every two days. The MODIS land (MODLAND) product suite range in spatial resolution from 250m to 1km. Several products are specifically related to biophysical parameters (Vegetation Index, LAI/FPAR, and NPP) (Justice et al., 2002). To avoid cloud cover and spurious effects for viewing and illumination angles (Wolfe, 2002) multi-day composites are made for each of the MODLAND products. Because of the daily (or two-day) repeat coverage, the multi-day composites nearly always contain some cloud free, useful, data. This results in a steady time-series of MODIS biophysical products. Further, these data are free and the lower spatial resolution implies reduced data volume.

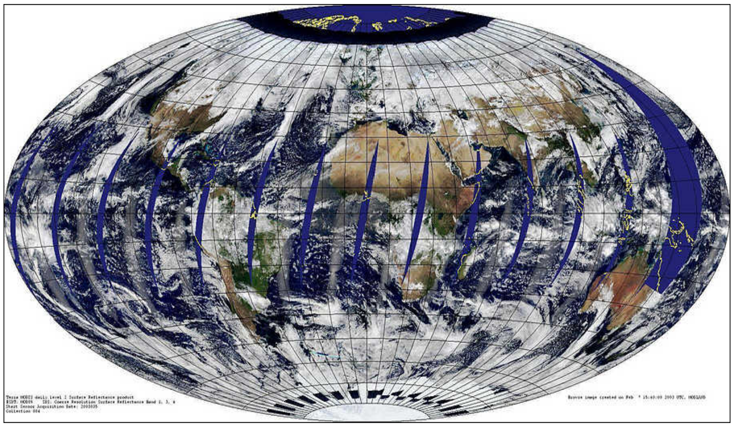


Figure 5: MODIS daily coverage map. In contrast to ETM+, the MODIS instrument provides daily coverage for areas beyond +/- 30^o and covers the entire glove every two days. The higher temporal resolution of MODIS biophysical parameters is particularly valuable for discriminating phenological patterns of rapidly proliferating exotic species, many of which derive invasive potential from subtle timing differences in the life history patterns between non-native and native species.

There are therefore compelling reasons to use MODIS data for invasive species research and applications. However, a potential drawback for the use of MODIS data is its lower spatial resolution. A primary research question – of particular and fundamental concern to the invasive species community – thus becomes:

> *Can we optimize the temporal resolution of MODIS data and the spatial resolution of ETM+ to improve geostatistical predictions of the location of invasive species and patterns in biodiversity?*

We believe the key to addressing this question lies in:

- Leveraging off existing, state-of-the-art, and well-quantified invasive species and biodiversity modeling efforts, which can provide a context for testing and implementing the MODIS/ETM+ time-series optimization effort;
- Building on a solid understanding of the MODIS land product suite and the associated quality assurance layers;
- Refining current methods of summarizing vegetation time-series data toward specific landscape characteristics;
- Integrating time-series summary values into the current modeling scenario.

Proposed Research

The goal of the proposed work is to provide USGS, the National Institute of Invasive Species Science, and formal and informal NIISS partners with new imagery products that summarize the time-series of MODIS vegetation products available since early 2000. The products will be tailored to the specific needs of the invasive species research and management communities. We also propose to create a data management capability that will allow integration of these summary products into the existing USGS invasive species modeling framework.

Objectives
The objectives of this project are as follows:

- Create a capability to extract time-series summary statistics from MODIS data to include within the ISFS to produce thematic, interdisciplinary data products for invasive species management and policy decision-making;
- Contribute to NASA's goal of broadening the use of EOS data by engaging a new community of users with an urgent, focused, national-scale data product need;
- Integrate these new data products into a coherent and comprehensive solution of national scope: the National Invasive Species Forecasting System;
- Expand and consolidate a broad coalition of users from across the invasive species research, education, and applications communities; and
- Make the resulting products broadly available on a sustainable basis through the USGS National Institute of Invasive Species Science and National Biological Information Infrastructure.

Approach and Rationale
Our overall approach will be to summarize vegetation time-series data for specific landscape characteristics that are known or presumed to have biological relevance to fundamental processes that drive exotic invasions. In a sense, we are attempting to bridge between a purely statistical approach and a more mechanistic understanding of invasive species. Our assessment of the usefulness of the MODIS summary values will thus require close collaboration with domain scientists, and we will work through an iterative process based on input from researchers and results from established statistical methodologies

Research Strategy
Over the three-year course of this project, we will consider a set of focus study areas, the use of all MODIS land products, and multiple methods to summaries the MODIS time series. For each focus study area we will consider any of the MODIS land products that might help predict the presence or spatial pattern for the invasive species of that area. We propose to comprehensively examine alternate characterizations of multi-year data sets appropriate for each focus area. Summary approaches will include DFT analyses, function fitting, and autoregressive modeling.

Focus study areas
We propose to analyze several data-rich landscapes in several states, i.e., "Intensive Sites," and the state of Colorado as an "Extensive Site." The intensive sites were selected to respond to the top research priorities of several land management agencies, to directly address targeted issues, and to focus on selected target species and species groups (Table 1).

Table 1: Intensive and extensive sites to be analyzed in this study.

Intensive Sites	**State**	**Primary Agency**	**Target Issue**	**Target Species (or groups)**
Rocky Mountain National Park (RMNP)	CO	National Park Service	Preserve native plant diversity in mountain parks	Patterns of non-native plant species
Cerro Grande Fire Site, Los Alamos (CGFS)	NM	USDA Forest Service	Vegetation restoration after wildfire in pine forests	Cover of native and non-native grass species
Grand Staircase-Escalante National Monument (GSENM)	UT	Bureau of Land Management	Grazing and plant invasions in arid ecosystems	Carbon storage of invasive cheatgrass and tamarisk
Konza Prairie, Long-Term Ecological Research Site (KP-LTER)	KS	LTER Program/ NASA	Estimating phenology and productivity in tallgrass prairie and detect woody encroachment	Rare tallgrass and herbaceous species
Extensive Site				
State of Colorado (CO)	CO	State of Colorado and State Weed Coordinator	Map sites vulnerable to plant invasions	Top 5 invasive plant species throughout the state

The Intensive study sites represent a broad spectrum of topography, ecosystem types, and climate regimes, maximizing the general applicability of the proposal. For example, in Rocky Mountain National Park, Colorado, 14 vegetation types were characterized in a 55,000-ha portion of the Park ranging from 2500 m to 3660 m elevation. These included lodgepole pine (*Pinus contorta*); aspen (*Populus tremuloides*); ponderosa pine (*Pinus ponderosa*); wet meadow (dominated by *Poa palustris, Deschampsia caespitosa*, and *Poa interior*); dry meadow (dominated by *Carex helianthus* and *Artemisia tridentata*); mixed conifer (may include *Abies lasiocarpa, Picea engelmanii, Pinus contorta, Pinus flexilis*); and alpine tundra. Most of the area was moderately grazed by elk and deer. Over 150 20 m x 50 m sample plots were randomly located in a 54,000-ha region of the Park (see Stohlgren et al 1997, 1999, Chong 2002) including many rare and common habitats. This site presents an opportunity to use time-series MODIS data and Landsat imagery to model the patterns of native and non-native plant diversity and cover in a popular National Park.

In the Cerro Grande/Los Alamos area of New Mexico, we have established 100 multi-scale vegetation plots in a 30,000-ha mixed conifer forests, burned in 1999 in a wildfire. The pine forests were extremely vulnerable to wildfire and subsequent erosion. Restoration activities included planting non-native grasses on steep and moderate slopes. This site provides an opportunity to use time-series MODIS data and Landsat imagery to model the effects of vegetation restoration and the invasion of non-native grasses after wildfire.

The Grand Staircase-Escalante National Monument is an arid-desert ecosystem on the Colorado Plateau in Utah. The Monument ranges in elevation from 1,370 m to 2,530 m and encompasses over 850,000 hectares. Over 350 20m x 50m multi-scale vegetation plots were selected in 18 vegetation types with the most common types being pinyon, pinyon-juniper, blackbrush, riparian, and desert grasslands, with a few rare types including ponderosa pine and aspen. Additional sites were randomly located in rare habitat types (e.g., wetlands, washes, and relict plant habitats as they were encountered in the field (Stohlgren et al. 2001, Bashkin et al. 2002). The presents an opportunity to use MODIS time-series and Landsat data to identify habitats vulnerable to invasion by cheatgrass (*Bromus tectorum*), which creates a severe wildfire hazard in the Monument, and *tamarix spp*., salt cedar, which steals valuable water from downstream consumers.

The Konza Prairie is a 3487-ha tallgrass prairie research site that has been part of the NSF-funded Long-term Ecological Research (LTER) network since 1981. Over 97% of the tallgrass prairie in the US has been lost to agriculture, urbanization, and fire exclusion (resulting in forest invasion). The Konza LTER program has documented some important changes in woody plant cover associated with long-term fire and grazing treatments (Knight et al. 1994, Briggs et al. 2002), and has begun assessing the causes and consequences of this ecosystem conversion from C_4 grass to C_3 woody plant dominance (Hoch et al. 2002). Using MODIS data to quantify the rate and extent of woody vegetation expansion into grasslands (both on Konza and in the surrounding Flint Hills area) is of keen interest for LTER researchers and would tie to other ongoing projects funded by NSF, the NASA LCLUC program (personal communication, John Blair), and EOS Land Validation (Morisette et al. 2002). Whole watersheds experiments with a range of long-term fire and grazing treatments (Knapp et al. 1998), combined with hundreds of vegetation sampling plots and long-term climate date provide a sound basis for evaluating the phenology and productivity in this ecosystem.

The Extensive Site, the state of Colorado, provides an excellent opportunity to evaluate issues of scaling from intensive sites (Rocky Mountain National Park) to much larger areas. Scaling information for landscapes to regions requires integrating moderate resolution remote sensing data (Landsat data), coarse-scale but high frequency MODIS data, multi-scale vegetation data at many sites, and improved computer storage and processing capabilities. The State of Colorado has asked our assistance in mapping the top priority invasive plant species throughout the state. Predictive spatial models of invading plant species have never attempted at state-wide scales. However, the ISFS is now compiling the requisite field and ancillary data for the state, and so in year 3 the timing will be right for applying MODIS' wide area coverage and summary techniques to this regional scale.

MODIS land products

For each study area we will subset all MODIS products that might be relevant for that area and its particular target issue. Morisette has worked with the MODIS adaptive processing system (MODAPS) to operationally subset MODIS products for specific areas. This system is available for extracting subsets of all MODIS land product for each focus study area. The command line version of the MODIS reprojection tool (http://edcdaac.usgs.gov/tools/modis/)

will be used to reproject and reformat the MODIS subsets into a projection and file format suitable for the Invasive Species Forecasting System (e.g. UTM and GeoTiff). IDL software (Research Systems, Inc, Boulder, CO) will then “stack” the data through time and run routines to summarize each pixel through time.

The particular MODIS products to be considered will depend on the target issue and our understanding of how that issue can manifest itself onto a given MODIS product. The specific MODIS land products that will be investigated for each site are listed in Table 2. For each site, we will utilize the MODIS surface reflectance bands and the BRDF/Albedo product for general radiative properties of the area. In addition, all biophysical products will be considered: VI, LAI/fPAR, and NPP. The snow, land surface temperature, and fire products will be considered for sites where they might provide insight into the location and distribution of the target issue. The land cover compositing period for the Land Cover and VCC/VCF products is so long it will not be necessary to summarize the time series of these products. However, these data will be subsetted and available for each study area for possible inclusion in the stepwise regression modeling. All data extracted for the proposed work will be ingested and available within the ISFS.

Table 2: MODIS products considered for each focus study area

Product Suite	**Product**	DAAC	RMNP	CGFS	GSENM	KP-LTER	CO
MOD09	Surface Reflectance	EDC	**X**	**X**	**X**	**X**	**X**
MOD11	Surface Temperature and Emissivity	EDC		**X**		**X**	**X**
MOD43	BRDF/Albedo	EDC	**X**	**X**	**X**	**X**	**X**
MOD10	Snow Cover	NSIDC	**X**			**X**	**X**
MOD29	Sea Ice Extent	NSIDC					
MOD13	Vegetation Indices	EDC	**X**	**X**	**X**	**X**	**X**
MOD15	Leaf Area Index/Fraction of Photosynthetically Active Radiation (LAI/FPAR)	EDC	**X**	**X**	**X**	**X**	**X**
MOD17	Net Primary Vegetation Production (NPP)	EDC	**X**	**X**	**X**	**X**	**X**
MOD12	Land Cover and Change	EDC	**X**	**X**	**X**	**X**	**X**
MOD14	Thermal Anomalies and Fire	EDC		**X**		**X**	**X**
MOD44	Vegetation Cover Conversion/ Continuous Fields (VCC/VCF)	EDC	**X**	**X**	**X**	**X**	**X**

Additional details available:
For Land Processes Distributed Active Archive Center (LP DAAC)
http://edcdaac.usgs.gov/modis/dataprod.html
For National Snow and Ice Data Center (NSIDC DAAC)
http://nsidc.org/

Time-Series Summary Methodology

Our efforts to summarize the MODIS time-series will involve three types of analyses. We will examine DFT approaches, function fitting, and autoregressive techniques. In our preliminary results (discussed below) we have examined the annual component of the DFT, but we will also examine other components such as biannual cycles and higher frequencies introduced by sudden events such as fires. We will test for inter-annual variability by examining the residuals of a constant amplitude yearly sinusoid, but we will also explore windowing and transforming each year of data separately. We will examine whether explicit fits to the time-series data will generate better characterizations of the information content for individual sites and target issues. We will begin by implementing the asymmetric Gaussian fitting of Jönsson and Eklundh (2002) but will explore alternate functional forms that are suggested by our data exploration. We will also explore autoregressive time-series models (e.g. ARIMA, Chapter 13 Venables and
Ripley, 1999). For these time series analyses we have experience with, and availability of, both IDL and S-PLUS (Mathsoft, Inc., Seattle, WA) software.

Our use of the extensive MODIS product suite will be founded in Morisette's five years of experience as the MODIS land team validation coordinator, with particular attention to the quality assurance and "known issues" for each product (http://landdb1.nascom.nasa.gov/QA_WWW/). Stohlgren and Schnase's ecological expertise will guide the physical understanding of the target issue and how best the multiple MODIS products can be summarized to capture a signal related to that issue. Pedelty's extensive experience with signal processing and IDL code will be used to implement various summary methods.

It is crucial to stress the importance and significance of conducting the proposed work within the existing NIISS framework. Our assessment of the usefulness of the MODIS time-series summary values will be an iterative process, based on both the input from researchers and results from the established statistical methodologies. Input from the researcher familiar with a particular site and species can help determine how landscape dynamics of interest will manifest themselves onto the MODIS data through time. These different phenomena imply the need for different methods to summarize the MODIS time-series. In turn, various methods can be statistically evaluated through the regression modeling; resulting in an empirical evaluation of the usefulness of a given summary value. The predicted surface and residual/error analysis will point to patterns potentially missed by the model and the current summary method. Returning to the biology of the system, different time-series summaries of various MODIS products can be explored.

Preliminary Results

To demonstrate the utility of incorporating MODIS time-series data into the USGS modeling approach, we compiled two years worth of MODIS data over three of the focus areas. The MODIS data used were 8-day composite, 250m resolution normalized difference vegetation index (NDVI) images created from the MOD09 surface reflectance product. The study areas are the Cerro Grande Fire Site, Grand Staircase-Escalante National Monument, and the Rocky Mountain National Park.

We created a set of prototype IDL procedures to assemble and characterize the NDVI time-series at each of these sites. We read the NDVI values for each date separately, and then stacked the complete set of images to build a cube with time as the 3rd dimension. We also read and interpreted the critical quality assurance (QA) bits, which we used to screen out unreliable composite pixels at each date, e.g. those that still contained clouds. We then apply a low-pass filter to generate the final time-series for each pixel, currently an 8-sample boxcar smoothing.

In principle we could then use the entire NDVI time-series as additional independent or predictor variables in a very large regression, but we have instead focused on deriving summary information for each pixel. We used the Discrete Fourier Transform (DFT) to extract this information (Moody and Johnson, 2001). Currently we retain the average of the time-series, which is the first (or DC) term of the Fourier transform. We also retain the real and imaginary parts of the annual frequency component, from which we determine the magnitude [= sqrt(real^2 + imaginary^2)], and phase [= tan^{-1} (imaginary / real)]. In effect we are using the DFT to find the best-fit yearly sine wave at each pixel, which is characterized by the average NDVI, the peak-to-peak range of the NDVI, and the date of peak NDVI. These three summary values are depicted for one pixel's time series in Figure 6 (left).

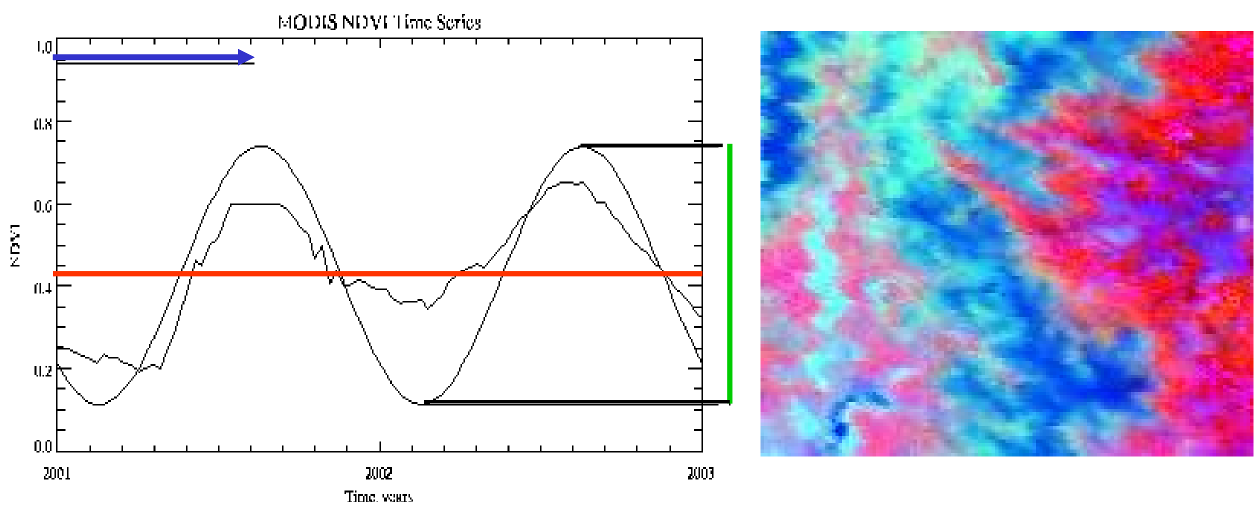
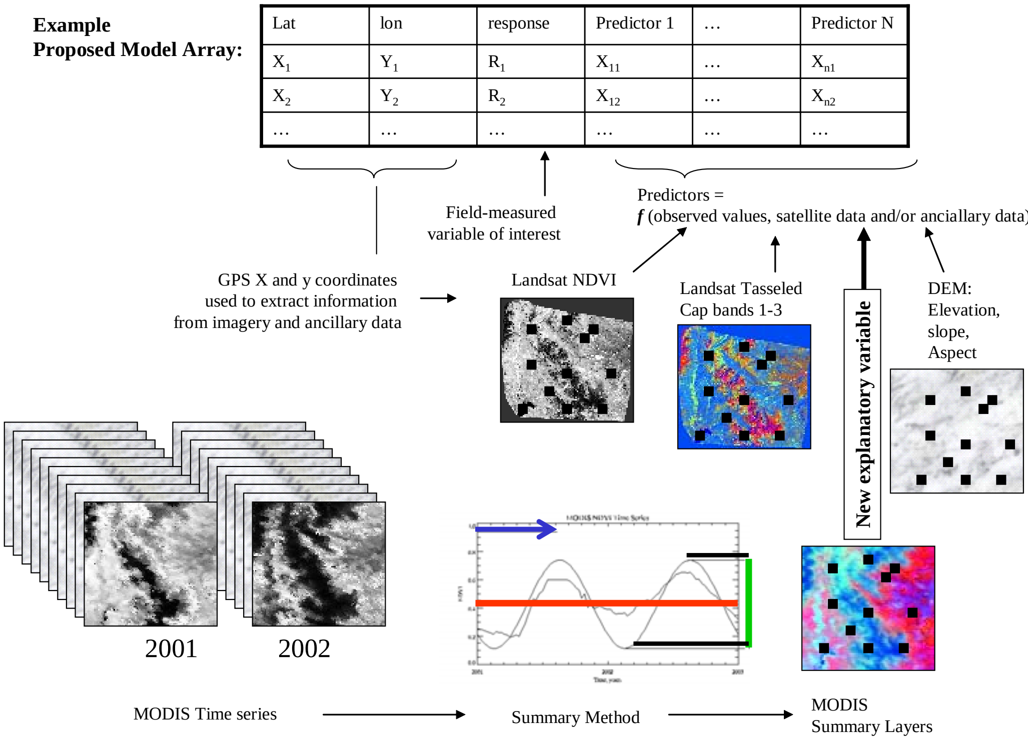


Figure 6: Example DFT summary for one pixel (left) and entire RMNP area (right) over two years:
red = average NDVI
green = peak-to-peak range of the NDVI
blue = date of peak NDVI

These summary values were extracted for each MODIS pixel covering each study area. The three summary values shown in Figure 6 (left) were then used to generate the false color image with corresponding red, green, and blue bands shown in Figure 6 (right). From this image we extracted the summary values for each field location and incorporated them into the modeling framework as three additional predictor variables (Figure 7, in contrast to Figure 3).

We then ran stepwise linear regression models with and without the MODIS data. In Table 3 we show the variability explained by the models at each site, as measured by R^2. In two of the three cases, the variability explained by the models that included the MODIS variables was significantly higher than the models without the MODIS variables. And in all three cases, the final model resulting from the stepwise regression included one or more of the MODIS time-series variables. While these results are limited to one MODIS product and one method of summarizing that product's time series, these initial results provide solid evidence that summaries



values from the MODIS time-series can improve the predictive capacity of the linear regression models.

Figure 7: MODIS time-series summaries have been incorporated into the USGS modeling framework. Preliminary results show significant improvements in R^2 values over what had been obtained early using ETM+ values alone.

Table 3: R^2 values for experimental model runs on three study sites with and without the use of MODIS time-series summaries.		
Study Site	R^2 without MODIS	R^2 with MODIS
Rocky Mountain National Park	0.32	0.41
Grand Staircase Escalante National Monument	0.19	0.24
Cerro Grande	0.36	0.36

Advantages and Impacts

The proposed MODIS time-series summary products will make a direct contribution to the joint NASA/USGS decision support capability currently being developed through the Invasive Species Forecasting System, as highlighted on the NASA Roadmap for Invasive Species (Figure 8). This proposed work makes a direct contribution to the paths shown as yellow text on the blue background. Work at the extensive study area in year three of the proposed work, covering the entire state of Colorado, will provide pathfinding activity for the text shown in red.

The joint NASA/USGS effort to develop the ISFS began in 2001. As a result, in part, of the NASA collaboration, USGS formed the National Institute of Invasive Species Science in 2002. The Institute is housed in the USGS Fort Collins Science Center (formerly Midcontinent Ecological Research Center) in Fort Collins, Colorado, and was dedicated in August, 2002. The Institute is the physical embodiment of USGS's invasive species decision support capability.

NASA is working closely with USGS leadership to coordinate these joint activities, including USGS Director Dr. Chip Groat, Associate Chief Biologist and NBII Director Gladys Cotter, and Doug Polson, Director of the USGS Fort Collins Science Center. USGS is committed to increasing investment in invasive species research and applications and to providing long-term support for the operational capabilities of the ISFS.

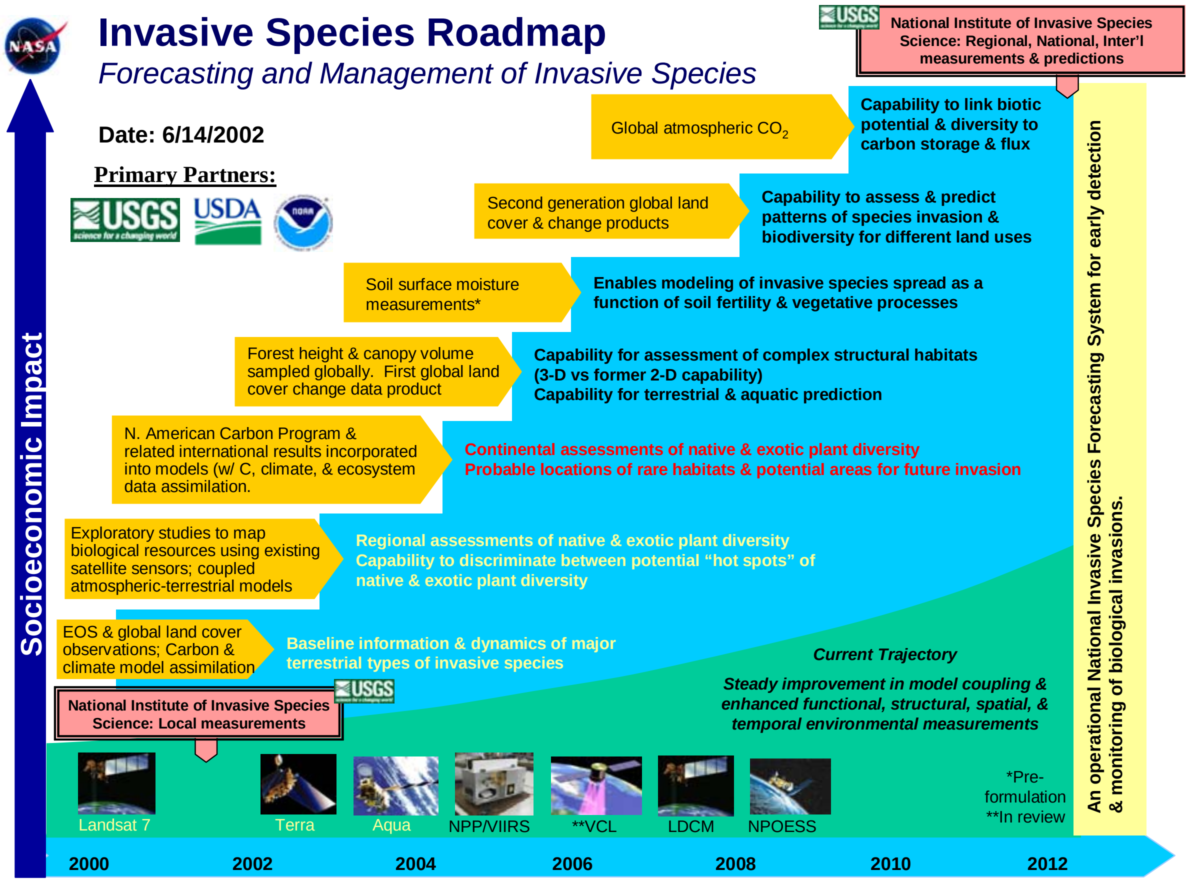


Figure 8: NASA's Invasive Species Roadmap

Project Management

Personnel

Our Investigator Team is broadly experienced and in an excellent position to assure a successful result. Dr. Morisette is a research scientist in NASA Goddard's Laboratory for Terrestrial Physics. His research is focused on the application of multi-resolution satellite imagery to ecological processes, to which he brings a strong background in statistics and geospatial statistical modeling. Dr. Morisette is coordinating the MODIS land team validation activities and is also chair for the Land Product Validation Subgroup of the CEOS Working Group on Calibration and Validation, which uniquely positions him to combine his detailed knowledge of the MODIS suite of data products with his familiarity with the long-term study sites.

Dr. Pedelty is a research scientist in NASA Goddard's Biospheric Sciences Branch. He has a broad and diverse background in physics and astronomy, with a particular interest in high performance computing and image/signal processing. He is former Principal Investigator in NASA's Computational Technologies (formerly High-Performance Computing and Communications) program and, as Co-Investigator on the current Invasive Species Forecasting System work, is leading our development of parallel geostatistical algorithms and code. He has worked in Earth remote sensing for the last decade, and was critical for the pre-launch testing of the ETM+ instrument on Landsat 7.

Dr. Stohlgren is the Director of the National Institute of Invasive Species Science at the US Geological Survey's Fort Collins Science Center. He has over 20 years experience in the Department of Interior and is author of over 100 scientific publications on assessing landscape-scale and regional effects of land use change on natural ecosystems; monitoring the effects of multiple stresses on native and exotic plant diversity; linking information at landscape, regional, and national scales; and developing GIS-based, predictive spatial and temporal ecological models to guide the management of public lands. He is on Task Forces of the National Invasive Specie Council, a Senior Scientist at the Natural Resource Ecology Laboratory and teaches courses in ecology at Colorado State University.

Dr. Schnase is a Senior Computer Scientist in NASA Goddard's Earth and Space Data Computing Division. He has worked extensively with the biological community to build capabilities in biodiversity and ecosystem informatics. He has served as chair of the bioinformatics working group of the President's Committee of Advisors on Science and Technology Panel on Biodiversity and Ecosystems and is a member of the OSTP/CENR Subcommittee on Ecological Science. Dr. Schnase has been involved in community information system development for many years and is the Principal Investigator on the NASA-funded invasive species activities that form the basis for the current proposal. Dr. Schnase is a Fellow of the American Association for the Advancement of Science and Co-Chair of the Science Advisory Committee of the National Research Council / National Biological Information Infrastructure World Data Center for Biodiversity and Terrestrial Ecology.

Responsibilities

Dr. Morisette will serve as Principal Investigator and will have overall responsibility for project management. He will work closely with Dr. Pedelty to design, implement, and test prototype algorithms to characterize the MODIS time-series data. In consultation with Dr. Stohlgren, Morisette will provide the statistical expertise to evaluate the utility of different time-series characterization for invasive species spatial predictions. Morisette and Pedelty will work with a half-time programmer/analyst contractor to acquire MODIS and precipitation data sets for the selected field sites. The programmer/analyst will manage the data sets, clean up and document the prototype code, and generate operational time-series characterization products. Stohlgren will also coordinate the selection of the study sites. Dr. Schnase will provide high-level scientific and technical guidance on integration of time-series products into the Invasive Specie Forecasting System and will work with Stohlgren to facilitate broader use of these data within the invasive species research and management communities.

Schedule

In the first year of the study, we will concentrate on Rocky Mountain National Park, Colorado, and the Cerro Grande/Los Alamos area of New Mexico. For these sites we will consider relevant individual MODIS products. In the second year of the study, we will consider combinations of MODIS products for these two sites, e.g. the relative timing of green-up after snow-melt in RMNP and vegetation signal before and after fire events. Also in year 2 we will include the Grand Staircase-Escalante National Monument, predicting the spatial distribution of tamarisk and cheatgrass. In year 3 of the study, we will apply the ISFS framework to model woody encroachment in the Konza Prairie region. We will also scale up to map invasive species for the entire state of Colorado. We will use the experience gained in year 2 on combining MODIS products to evaluate combinations appropriate for these two sites. Throughout the two years we will apply spectral (e.g. DFT), function fitting, and autoregressive (e.g. ARIMA) time-series models. The resulting model parameters will serve as the time-series summary statistics.

Expected Results

The primary product of this work will be a series of MODIS time-series summary data products specifically designed to allow scientists and resource managers to model and analyze regional-scale biotic resources. The information products will be formatted for ingest by predictive spatial statistical models that produce electronic and printed maps of potential “hot spots” of native plant diversity, including: (1) probable locations of rare habitats, (2) probable locations of relict/unique species assemblages, (3) potential areas of future invasion, (4) spatial auto-correlations with cross-correlation statistics for single exotic species, (5) accuracy assessments of native and exotic plant diversity, (6) evaluation levels of uncertainty in maps of natural resources, and (7) classification and regression trees for map accuracy using multi-phase (double) sampling. These products are critical elements of an invasive species DSS.

Scientific and technical results will be communicated to the user community through a series of annual workshops and presented at national meetings (ESA, SAF, ASPRS, etc.). The system, software, documentation, and information products will be delivered to the user community via the USGS National Institute of Invasive Species Science and the National Biological Information Infrastructure website at *www.invasivespecies.gov*.

Conclusion

The National Invasive Species Council has noted that “no comprehensive national system is in place for detecting and responding to incipient invasions.” Yet the threat of invasive species is perhaps our most urgent economic and conservation challenge. The National Invasive Species Forecasting System is an application of extraordinary promise for managing one of the 21st century’s most formidable threats of natural disaster. NASA’s unique national role in developing remote sensing capabilities and in understanding the Earth system, coupled with the urgency of the invasive species threat, the ready partnership of the USGS, and the compelling societal and economic benefits of this effort, make broader use of MODIS data – and, in particular, broader use of high temporal resolution MODIS data -- a critical next step that is strategically important to both NASA and the Nation.